REMARKS

I. Status:

Applicant thanks the Examiner for his thoughtful review of the application, as set forth in the Office Action dated May 6, 2009. Due to the finality of the Action, and the addition of claims herewith, Applicant responds with a Request for Continued Examination. The within submission, in support of the request, addresses all grounds of rejection in the May 6 Office Action, and provides an explanation of the added claims.

Claims 7-59 are in the application. Claims 7-16 were in the application prior to the present amendment; this amendment adds claims 17-59.

In the Action, the Examiner stated that he had considered Applicant's response to the previous rejection under 35 USC § 102(e) based on Gringeri et al., U.S. Pat. No. 6,233,226 ("Gringeri"), but that the issue of anticipation by Gringeri was rendered moot by new grounds of rejection. The new rejection was a rejection under 35 USC § 103(a), based on a combination of Gringeri with Baker, U.S. Pat. No. 6,449,719 ("Baker"). Claims 7-16, which were all of the claims in the application at the time, were rejected on this basis. The Action was made final.

Applicant respectfully traverses the rejection under § 103(a).

In addition, Applicant amends the Specification, and adds additional claims, as noted above, which Applicant submits should be allowable, if the Examiner agrees that the rejection under § 103(a) has been overcome.

Reconsideration of the application in light of the amendments to the claims and specification, and the following remarks, is respectfully requested.

II. <u>Discussion of Present Application:</u>

Applicant's invention, as delineated by claims 7-16 (as amended) and claims 17-59, provides in various aspects a method and system for distributing streaming media data over a network to one or more end users (clients). As a result of the buffering methodology employed, the system and method virtually eliminates disruptions in the playback of the media for the users arising from transient network congestion that afflicted previous distribution systems. Using previously available methods, it was highly likely that during distribution of media data, the vagaries of network data traffic would result in interruptions of the continuous transmission of the data (*see* Par. [0007] of the published specification). If the interruption were longer than the interval required to replay whatever data was cached in the client computer (or like device), the continuity of playback would be disrupted. Such disruptions are frequently referred to as "dropouts" (Par. [0008]). Methods for eliminating dropouts are highly sought in the networking arts, to avoid the user inconvenience and frustration that results from gaps and repeated delays in content delivery.

Applicant's invention is able in the steady state to send data elements from the server to the user's playback device at the continuous playback rate of the streaming media (Par. [0021]). In addition to virtually eliminating interruptions, the present method and system also permits the playback of media data to begin nearly instantaneously after a user's request (Par. [0032]). By contrast, previous systems necessarily entailed a significant delay to permit downloading of ex-

tended amounts of data (displaying a message such as "Buffering"), before playback could reliably be initiated (Par. [0008]).

The present application presumes an Internet Protocol (IP) networking context (see Pars. [0002] and [0032]). However, it does *not* presume a Quality of Service (QoS) guarantee: "Internet connection quality can vary rapidly over time." Par. [0007]. In various aspects, the present invention, as reflected in the present set of claims, is designed to overcome the *lack* of a QoS guarantee in an ordinary Internet connection. The claimed invention addresses the following problem: How to provide a substantially continuous playback at the user end of media that plays at a given rate, over a connection that may fall well below that rate for significant periods that would be intolerable if apparent in the playback stream. The solution reflected in the present claims is to provide an uninterrupted data flow at the user end by buffering at the application level (OSI layer 7) (or at least at a level higher than the network transport layer (OSI layer 4)), coupled with a variable-speed transport from the server.

In one aspect, the problem may be addressed through a buffer associated with the server. The server-side server buffer may be used as follows. The server receives media data from a media source, such as a broadcast station, or alternatively, from another source, such as a disk file (Par. [0041]). However, the server does not relay this data in real time to the users' media player. Rather, the server loads this data into the server buffer until it contains at least a desired amount of data, e.g., about a minute's worth (Par. [0042]). Once the server buffer is sufficiently full, and after the user's media player connects, the contents of the server buffer are sent (in playout order, as would be the case, for example, in a FIFO

buffer) to the user's media player at a rate faster than the real-time playback rate. Alternately, where, for example, pre-recorded media is available from a local disk file, the data may be read into a first buffer, and a moving data window within that buffer may be used to similar effect (Par. [0041]), to send an initial supply of data at a rate faster than the playback rate. The rate at which data is sent to initially fill the user's player's buffer may be as fast as the data rate of the user's connection will allow (e.g., Par. [0044]), but in any case, the data will be sent more rapidly than it is played out by the user system (Par. [0032]). The transmitted media data is stored in the user's media player buffer, where it is then decoded and played out at the appropriate real-time playback rate (Par. [0044].

The fact that the player is continuously playing at the play rate of the media means that after the user player's buffer has been filled, the server can (indeed, should) cut back its send rate and continue sending media data at about the playback rate. If, due to a network interruption or slow-down, the user player's buffer depletes below full, the server can detect this, e.g., through built-in facilities of the transport protocol (e.g., TCP) (Par. [0022]). Data will back up in the server buffer during the interruption or slowdown (Par. [0021]) (or, in a "data window" implementation, the data window will remain stationary). When it is detected that the interruption or slowdown has cleared, the server can temporarily step the transmission rate up in order to top up the user player's buffer (Par. [0022]). This rapid recovery feature enables a user's media player to survive most dropouts and deliver an excellent experience, even under degraded network conditions. As explained in the specification, in a scenario as described above, there would always be approximately 30 seconds (or whatever minimum interval

is designated) of buffer material between the server buffer and the user buffer – though the allocation between these two buffers may slide back and forth during the course of media delivery (Par. [0047]).

In effect, the above-described mechanism creates an elastic system of buffers between the server and the player. Because of the relatively fast initial fill rate of the player buffer, the player can start playback almost immediately—as soon as only a few media data packets have been received. This virtually eliminates waiting at startup.

In addition, a single server in the present system can service a plurality of players (Par. [0027]). Since every player might be at a slightly different point in the stream (due to different times of individual connection and individual circumstances of network latency, lost packets, etc.), the present process needs the capability of keeping track of each player individually. The TCP/IP protocol provides the means to do this for every connection port, coupled with a set of pointers tracking in the server buffer the data elements last sent to the user. Alternatively, instead of server-side pointers, progress may be tracked on the user side by identifiers associated with the media data elements, and the identifier for the last element received can be communicated back the server via TCP when the user system requests more data, so the the server may send the next sequential media data elements (Par. [0029]).

Discussion of Specification Amendments

The cross-reference to related applications has been amended by inserting a caption (i.e., "CROSS-REFERENCE TO RELATED APPLICATIONS"), which previously had been missing.

Language has also been added specifically claiming of the benefit of the filing date of U.S. Provisional Application Serial No. 60/231,997. U.S. Application Serial No. 09/819,337, parent of the present application, claimed the benefit of the filing date of the provisional application. The inventor's declaration dated March 24, 2001 (the parent case declaration), which is filed in this case, expressly claimed the benefit of the prior provisional application. Hence, the added cross-reference does not incorporate new matter.

The amendment also corrects the typographical error in the serial number of the parent case. The proper serial number of the parent case was correctly noted in the original transmittal, and is currently correctly noted in the electronic file wrapper. Hence, this correction does not add new matter.

In addition, for purposes of consistency and clarity, the specification has been amended to correct an inadvertent omission. During the filing of the present continuation application, the final two paragraphs of the parent application (US Application Serial No. 09/819,337), and part of the third-to-last paragraph were inadvertently not included, and are hereby restored. It is respectfully noted that the parent application was expressly incorporated by reference in its entirety in the present application, as set forth in Par. [0001]. Accordingly, no new matter has been added by the incorporation of these omitted paragraphs, which were included in the parent application as filed.

Discussion of Claim Amendments

In order to emphasize the patentable distinctions of applicant's contribution to the art, base claims 7 and 15 have been amended. In particular, these claims were amended to call for the sending of data elements from the server to a buffer in the user's computer at different rates, depending on the status of the user buffer. Initially, and whenever the user buffer is not full, the data are sent at a rate faster than the playback rate (the rate at which the data elements are played out). When the user buffer is full, the elements are sent at the playback rate. Accordingly, the user buffer is maintained at approximately a full level at all times, absent extended transmission disruption, yet playback from the user buffer ordinarily continues without apparent interruption.

New claims 17-59 have been added to provide adequate coverage for applicant's contribution to the art.

New independent claim 17 is directed to a method for distributing streaming media to at least one user system of at least one user. Like the method of claim 15, claim 17 calls for initially sending the streaming media to fill the user system's buffer at a rate more rapid than the rate at which the streaming media is played out by the user system. Thereafter, the method provides for detecting whether or not the user buffer is full. If it is, then subsequent sequential media data elements are sent to the user at about the playback rate. If the user buffer is not full, the next elements are sent at a rate more rapid than the playback rate. Claims 18-31 depend from claim 17 and recite additional features.

New independent claim 32 is directed to a streaming media server providing a buffering system for distributing streaming media to at least one user system of at least one user. Claims 33-45 depend from claim 32.

New independent claim 46 recites a machine-readable medium on which there has been recorded a computer program for use in operating a server for distributing streaming media via a data communications medium such as the Internet to at least one user system of at least one user. New claims 47-59 depend from claim 46.

Support for the foregoing amendments and the newly added claims is clearly provided by the original specification; with reference to the published specification, this support may be found, particularly, at Pars. [0002], [0007], [0008], [0021], [0022], [0027], [0029], [0032], [0041], [0042], [0044], and [0047]. Consequently, no new matter has been added.

III. <u>Art Rejections:</u>

<u>Claim Rejections under 35 U.S.C. § 103(a)</u>

Claims 7-16 stand rejected under 35 U.S.C. § 103(a) as being obvious over Gringeri et al. (U.S. Patent No. 6,233,226) ("Gringeri") in view of Baker, U.S. Pat. No. 6,449,719 ("Baker"). Applicant respectfully traverses this rejection.

35 U.S.C. 103(a) provides in pertinent part:

> A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. (Emphasis added.)

The United States Supreme Court, in *KSR Int'l Co. v. Teleflex, Inc.*, 550 U.S. 398, 82 USPQ2d 1385 (2007), ruled that the framework for determining obviousness under § 103 remains the one stated in *Graham v. John Deere*, 383 U.S. 1, 148 USPQ 459 (1966). According to the *Graham* framework, obviousness is a question of law that is determined based on the results of the following three underlying factual inquiries (sometimes referred to as the "*Graham* inquiries"):

(a) determining the scope and content of the prior art;

(b) ascertaining the differences between the claimed invention and the prior art; and

(c) resolving the level of ordinary skill in the pertinent art.

In order to reject a claim for obviousness, the Examiner must make factual findings that are sufficient to support the ultimate legal conclusion of obviousness. According to the MPEP, the Examiner must make "findings of fact concerning the state of the art and the teachings of the references applied", and "must provide an explanation to support [the] obviousness rejection" (MPEP 2141). As stated by the Supreme Court, "there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." KSR, 550 U.S. at 418, 82 USPQ2d at 1396. The MPEP further elaborates, stating that "[t]he focus when making a determination of obviousness should be on what a person of ordinary skill in the pertinent art would have known at the time of the invention, ***and on what such a person would reasonably been expected to have been able to do in view of that knowledge***." MPEP 2141, emphasis added.

The MPEP goes on to state: "[o]nce Office personnel have established the *Graham* factual findings and concluded that the claimed invention would have been obvious, the burden then shifts to the applicant to (A) show that the Office erred in these findings, or (B) provide other evidence to show that the claimed subject matter would have been nonobvious." MPEP 2141(IV).

Accordingly, we turn now to a review of the *Graham* analysis, as applied to the claimed invention.

(a) First *Graham* Inquiry: Scope and Content of the Prior Art.

(1) Gringeri

Gringeri concerns a system and method for analyzing and transmitting video streams over a switched network. Gringeri specifically discloses a system and method for scheduling and transmitting video streams over an Asynchronous Transfer Mode (ATM) network. Col. 1, lines 22-25. In ATM communications, messages are broken down into "cells" having a fixed length of 53 bytes. Dividing the information into cells, each having a fixed length, allows the information to be transported in a predictable manner. Unlike IP networks, where there is essentially free-form routing over open-ended topologies without constraining the order or time of delivery of individual packets, the ordering and staged continuous delivery of cells is built into an ATM network at a low level.

Gringeri teaches that "A unique feature of ATM is the ability to provide guaranteed throughput levels." Col. 3, lines 10-11. The disclosure of Gringeri is directed at two particular ATM modes, called Constant Bit Rate (CBR) and Variable Bit Rate (VBR), which provide these guaranteed throughput levels. Each of

these modes allow the sender of a transmission to specify a number of parameters, by “contracting” for them in advance of the transmission. One such parameter is the Sustained Cell Rate (SCR), which is the calculated average allowable, long-term cell transfer rate characteristic of a specific connection. Since, in an ATM environment, the customer gets only the capacity (i.e., QoS) paid for, the objective of a provider of streaming media would be to contract for an SCR approximating (and certainly not greatly exceeding) the playback rate of the media. In the context of streaming media, SCR in an ATM environment will be comparable to the average playback rate of the media. Credits or debits from the contracted-for SCR rate are reflected in "tokens", which can be earned (for example by slack activity) or used (for example, by increased activity) during the course of communications. In particular, accumulated tokens may be used during transmission to cover transmission of specified cells at the Peak Cell Rate (PCR), should this become necessary during transmission.

The problem addressed by Gringeri is how to distribute a video over a transport mechanism, such as that provided by an Asynchronous Transfer Mode (ATM) network, in which, according to the network protocol, the distributor will have to contract for a “sustained” or average throughput rate specified in advance, but wherein the video will have data transmission demands from time-to-time far in excess of the average transmission rate, and where there will only be a limited ability to send cells at the higher PCR rate.

Video of the type contemplated for distribution by Gringeri has a variable data rate. For example, MPEG2 encoded video, which is consumed by a viewer at a constant frame rate, has varying amounts of data associated with the individual

frames, depending on factors such as scene motion, etc. Still scenes may have low amounts of data per video frame, while high-action scenes may have much higher amounts of data per frame. Therefore, to supply data at a constant viewing rate (frame rate), it is necessary to be able to accommodate a potentially large range of data transmission rates. Therein lies a risk of starving (underflowing), or conversely overflowing, the user's receive buffer, resulting in interruptions and/or lost frames.

Gringeri, in addressing the distribution of such video, reviewed four types of ATM services (CBR, VBR, UBR and ABR), finding that only CBR (Constant Bit Rate) and VBR (Variable Bit Rate) ATM services provided workable bounds for maximum delay and delay variation (see col. 2, line 36 - col. 3, line 8), deemed necessary by Gringeri in order to provide a workable solution (and rejecting UBR and ABR for these purposes).

However, while CBR and VBR were the only ATM services deemed suitable by Gringeri to practice the distribution methods devised, neither of those protocols make any provision for "feedback" to the source distribution server about the state (fullness) of the user receive buffer.

Accordingly, the approach taken by Gringeri, to avoid user buffer overflow or underflow in the absence of direct feedback concerning the state of a user's buffer, was to "model" the user's buffer on the server side, and to use the results of the modeling to determine how much bandwidth to contract for, for a given video. There were two phases to this modeling, referred to in Gringeri as (1) the "pre-transmission video analysis phase" and (2) the "transmission phase". In the analysis phase, a model of the user's receive buffer was used in a preprocess step,

to predict what Sustained Cell Rate (SCR) is needed, and what peak cell rate (PCR) and maximum burst rate (MBS) will be necessary in connection with the selected SCR, to guarantee no underflows or overflows during transmission. Col. 8, lines 37-60. The selected SCR is effectively the average anticipated playback data rate of the video. In addition, the model was also run during actual transmission. Col. 9, lines 8-19. While the analysis phase determined the SCR, PCR and MBS to contract for, such as to guarantee that the video *could* be transmitted without interruption, it did not generate a time line for regulating the transmission, i.e., when during transmission to stop and wait or alternatively, to step the speed up to the PCR. In order to determine these actions in real time, the buffer model was also run during the "transmission phase", that is, during actual transmission, to act as a real-time control on the transmission process, using the SCR, PCR and MBS previously selected according to the analysis phase.

Actual transmission begins after the SCR, PCR and MBS parameters are set. The user buffer is first filled **at the SCR**—as discussed above, the SCR is roughly the playback rate. The user buffer is monitored during transmission by performing an internal server-side **simulation** comparable to the pre-transmission analysis. If the user buffer becomes full, transmission is **stopped** until a frame is removed from the user buffer. Thereafter, transmission is resumed at the SCR, but if tokens are available at that point, the transmission rate is increased up to the PCR until they are used up, and otherwise the cell is sent at the SCR. *See* col. 9, lines 8-19; col. 23, lines 30-35. Thus, while Gringeri provides for temporary speedups and slowdowns during transmission, they are based on modeling predictions, not actual feedback.

(2) Baker

Baker concerns an encryption mechanism to "copy protect" streaming media to restrict distribution to authorized viewers and to prevent transmitted streams from being recorded and retransmitted to others. Baker discloses the use of a UDP (User Datagram Protocol)/IP protocol. UDP/IP, unlike, for example, the TCP/IP protocol, sends data packets without checking up on whether the packets were successfully received, or received in the proper order. Thus, freed of these responsibilities, UDP/IP can function with very low networking overhead. In an UDP-based system, it is up to the software in the application layer to monitor and/or moderate data flow to the extent needed. Baker provides for the flow control aspects of the transmission by providing a "client flow control module 230 to obtain feedback from the data buffer in the client 710 and control the rate of the data stream to keep the client buffer as full as possible". If this module determines at any time that the client cannot accept any more data, the client flow control module will act to slow down or pause the data stream. Col. 8, lines 6-24. This is easily done in the context of Baker, because UDP is a connectionless protocol that does not establish a dedicated end-to-end connection, and thus can be turned on or off at will on the server side. The focus of Baker is on encryption and content protection. While Baker addresses flow control, it does not address the question of how throughput may be maintained over connections of variable quality without pauses, dropouts, lost frames, etc.

(b) Second *Graham* Inquiry: Ascertaining the Differences Between the Claimed Invention and the Prior Art.

There is a fundamental difference between the communications protocols that provide the respective contexts for Gringeri and the claimed invention. Unlike the present invention, Gringeri presumes the presence of a transport mechanism (i.e., ATM/CBR or ATM/VBR) that is not only reliable in terms of ultimate delivery, but provides guaranteed throughput and guaranteed bounds on delays. In the claimed invention, IP protocols are utilized, which do not provide these guarantees. Therefore, the claimed invention must implement additional methods, beyond the native capabilities of the network, to assure reliable content delivery. Furthermore, Gringeri, in contrast to the present invention, broadcasts transmissions to a plurality of users in lockstep, whereas the present invention discloses and claims mechanisms for individualized delivery of streaming media to users, based on the quality and state of each user's connection.

In the Internet and other IP-type networks, there is essentially free-form routing over open-ended topologies, without any constraint of the order or time of delivery of individual packets. IP packets can each take *different* routes from source to destination. By contrast, the ordering and staged continuous delivery of cells is built into an ATM network at a low level. See, e.g., background discussion at col. 2, lines 12-26 of Gringeri.

An ATM network could be thought of as a railroad with an endless series of rail cars to carry data – every car will reliably get to the destination in a set sequence. That's the only place it can go, and there isn't any interference from other rail cars – they're all in line, front to back, moving deterministically from one

end of the network to the other. In contrast, the Internet can be likened to a network of roads and highways with trucks, buses, and cars entering, exiting and switching from one road or highway to another in a chaotic fashion. Transit between two points can occur via many alternative routings. There are busy roads, collisions, roads and ramps that are broken, high-speed highways, low-speed roads, and "last mile" roads that sometimes barely operate (*e.g.* modems). *Internet routers are designed to throw away packets* if they get too busy, or if a link gets too busy. This is the inherent design, and it's up to the applications at the endpoints to figure out what to do about it. The Internet is a network in which a packet can travel from almost any endpoint to almost any other endpoint through a maze of interconnected routers – and in fact, the average packet traverses about twenty routers. The net result is that the Internet is clearly known to be both non-deterministic and unreliable.

Gringeri states that its methods can also be used in TCP/IP networks, but only insofar as "quality of service for a connection can be provided". Col. 23, line 67 - col. 24, line 1. "Quality of service" assurance means a **guarantee** of a certain level of performance to a data flow. While the ATM protocol does provide one level (UBR) with no guaranteed throughput levels, Gringeri deliberately does not use UBR. Rather, the techniques disclosed in Gringeri presume a quality of service (QoS) level having traffic descriptors with parameters including the peak cell rate, the sustained cell rate, and the maximum burst size. Col. 5, lines 30-31. At a given selected QoS level, there may be "jitter" on the order of milliseconds in the timing of the arrival of packets, but generally no major deviations from conti-

nuous sequential arrival. By contrast, the present invention, as claimed, does not presume any QoS level.

However, despite the potential advantages of guaranteed throughput, the ATM approach comes at a severe cost, because it demands the moment-to-moment availability of bandwidth adequate for the steady-state media playout rate for the requisite period of time. The TCP/IP protocol, on the other hand, permits a more opportunistic approach, in which continually varying amounts of bandwidth may be utilized.

The ATM transmission protocol as utilized by Gringeri inherently distributes a single program stream. Even if it were received by multiple users, each would receive *the same content in lockstep*. On the other hand, applicant's presently claimed Internet media method and system invention seeks to compensate for an unreliable network, while providing delivery of such media from either files or live broadcasting. It does so in a manner that inherently can be applied for either a single user or a large plurality of users, each having the ability to receive individualized content. As will be addressed, these fundamental differences between Gringeri and Applicant's claimed invention are material to the question of patentability.

One technique that is routinely encountered in IP protocol applications, but not generally employed over ATM networks is feedback-mediated flow control. Indeed, the Examiner recognized in the present action that Gringeri differs from the present claims in that very respect. As the Examiner correctly noted:

> "Gringeri does not explicitly teach [that the server's] transmission means is configured to receive notification from said user computer of the level of filling of said user buffer and to cause said server to cease sending said da-

> ta elements while said user buffer is full and thereafter to resume sending said data elements." (Action at page 3.)

Nor does Baker address the issue of buffering a connection over communications media of varying quality to ensure reliable, uninterrupted communications.

However, the Examiner proposed that Baker could be relied upon to supply the element of "client buffer feedback" found to be missing from Gringeri. The Examiner stated in this regard that:

> "Combining Gringeri with Baker would not require any significant modifications to Gringeri because the modeling taught by could be implemented independent of feedback with feedback providing an additional layer of client buffer protection." (Action at page 4.)

The Examiner here, while recognizing that Gringeri was specifically devised to work without client feedback, is suggesting that it would have been obvious to improve upon Gringeri's scheme by adding and additional layer of client buffer protection by superimposing on Gringeri a flow control functionality according to the feedback mechanism of Baker, and that it would not have required "any significant modifications to Gringeri" to have done so.

Applicant respectfully disagrees. Applicant submits that the combination suggested by the Examiner, to the extent that it even might be workable, would not have been obvious to one of ordinary skill in the art. The difficulty here is that Gringeri provides no means whereby an ordinarily skilled practitioner could engraft "client feedback" on the mechanism provided by Gringeri. That is because, in the ATM environment addressed by Gringeri, there is no suitable transmission protocol that has the capability of providing receive buffer feedback.

Fundamentally, Gringeri teaches away -- with good reason -- from the only ATM protocol that has the capability of providing client feedback -- "Available Bit Rate" (ABR). While ABR does provide feedback capability, as Gringeri itself points out, at col. 3, lines 3-5, and lines 9-10, that ABR does not provide bounds for the maximum delay and delay variation, thus rendering it unsuitable for the uses contemplated by the methods disclosed in Gringeri. In particular, ABR (like UBR, discussed above) lacks an SCR parameter (or any similar parameter), because ABR is incabable of guaranteeing a sustained cell rate. However, the ability to specify an SCR (as well as supporting PCR and MBS parameters) is critical to the modeling and data rate contracting necessary to implement the methods taught by Gringeri. *See, e.g.*, Fig. 5 of Gringeri (flowchart of analysis modeling phase) and discussion at col. 15, lines 17-30; and Fig. 10 (flowchart of transmission modeling phase) and discussion at col. 22, lines 16-21. These mechanisms are centered about determining an appropriate SCR, PCR and MBS for transmitting a given video program, and then trying to send the file at that SCR average speed. Because of ABR's inability to set such parameters, Gringeri cannot, consistent with its other teachings, use ABR, which is the one and only ATM protocol that would have allowed for client feedback.

(c) Third *Graham* Inquiry: Resolving the Level of Ordinary Skill in the Pertinent Art.

In the final step of the obviousness analysis, the Examiner must consider the level of ordinary skill in the pertinent art (and what type of practitioner has these skills), and, as instructed by the MPEP 2141 (supra), on "what such a person

would reasonably been expected to have been able to do in view of that knowledge."

Applicant submits that the level of ordinary skill in the art would be reflected in a telecommunications engineer with experience in network communications protocols including ATM and TCP/IP. Applicant submits that, for a variety of reasons to be discussed, the proposed modification of Gringeri, to the extent even desirable, would not be within the ordinarily expected capabilities of such a practitioner.

As discussed above, ATM protocols do not provide for feedback from the client side in any manner usable for reliable delivery of streaming video. Thus, to implement Baker-style feedback in the context of Gringeri would require going outside of the capabilities provided by ATM. That would require either a higher level network protocol that provided such functions (and there is no such high-level protocol for ATM), or application programming, on top of ATM, to set additional ATM channels between client and server specifically for flow control messages. Such an approach would then have to modify the disclosed mechanisms of Gringeri in some unspecified manner, in order to speed up or alternately stop transmission over the CBR or VBR channel as necessitated in accordance with the feedback -- and repeatedly re-adjust the transmission-phase buffer model as well, in order to reflect the changes that just occurred outside of that model, to keep the model in sync with the operative data transmission parameters. This is something that is not nearly as easily done as starting and stopping UDP transmissions as in Baker.

Further, in order for the added feedback mechanism suggested by the Examiner to constitute an "improvement", it must be presumed that that there are scenarios in which Gringeri's mechanism would fail. Gringeri flatly states that its buffer modeling method and "just ahead" planning mechanism "guarantee" no client buffer underflow or overflow during transmission. Col. 8, lines 59-60. We will assume, however, solely for purposes of evaluating the Examiner's proposed modification, that despite Gringeri's "guarantee" statement (which may well be correct), some video scenes could nevertheless require such sustained bursts of data that Gringeri's modeling projection would not provide for sufficient throughput at the contracted SCR rate to get through the entire scene, even after using the available number of PCR tokens (the fallback mechanism taught by Gringeri). On this assumption, such a development could (hypothetically) result in dropped frames.

Feedback could perhaps be provided in an ATM environment by adding a reverse ATM channel -- from client to server -- to send feedback messages. That in itself would ask a lot of the "ordinarily skilled practitioner" who was qualified as stated above. However, to make any use of the feedback, the underlying mechanism of Gringeri itself would have to be deeply reconfigured. A new ATM contract would have to be negotiated on the fly to kick up the transmission rate as needed to cover the burst, and then renegotiated again when the burst scenario ended. The "improvement" would have to calculate what increased rates to contract for, and when to back down to a lower, more economical, SCR rate. And it would have to reconfigure the running transmission phase buffer model to reflect these external changes, each time they occurred. As can be seen from Gringeri,

ther was considerable mathematical analysis necessary in order to calculate the proper SCR, PCR and MBS in the first instance. Adapting this detailed analysis to determining new rates in real time and readjusting buffer modeling on demand would be highly complex, to say the least.

In short, the "improvement" to Gringeri hypothesized by the Examiner would entail a wholesale redesign of Gringeri, which Applicant submits would be well beyond the capabilities of the ordinarily skilled telecommunications engineer. Furthermore, without knowing how to determine the new ATM parameters, how to time the transitions between different ATM contracts, and how to readjust the buffer model, which neither Gringeri nor Baker teach, there would be substantial uncertainty with regard to the expectations of success, a further factor that cuts strongly against a finding of obviousness. *See KSR*, 550 U.S. at 417, 82 USPQ2d at 1395; MPEP 2143.02 (entire section is captioned: "Reasonable Expectation of Success is Required").

Moreover, even if a skilled practitioner could implement such a scheme, it is not at all clear that it would improve over what could be done with the systems and methods that Gringeri already provides. While the Examner is not bound to accept Gringeri's claim to "guarantee" no buffer underflow or overflow, the Examiner must cite a good reason to doubt it, if suggesting an improvement to work around its assumed failure as a basis for finding obviousness. In this regard, the Examiner must recognize that, unlike either the present invention, or Baker, Gringeri presumes the presence of a data transport mechanism (i.e., ATM/CBR or ATM/VBR) with guaranteed bounds on throughput and delivery. Further, Gringeri takes substantial measures to avoid overflow or underflow events from

occurring during actual transmission, providing for a second round of user buffer modeling, in real time, during actual transmission (discussed at length at col. 22, line 6 to col. 23, line 54). Taken at face value, these factors would remove any advantage that would be provided by buffer monitoring via feedback. There is no indication that, in an ATM context, buffer modeling in response to the actual transmitted data, which is what Gringeri teaches as the second (transmission) phase of modeling, would be any less effective than feedback, in both sensing and predicting user buffer overflows or underflows. There is no reason to expect transmission-phase buffer modeling to be ineffective in the context addressed by Gringeri, unless the pre-transmission modeling that determined the SCR, PCR and MBS parameters in the first place simply does not work as claimed. As pointed out at the beginning of the present discussion of the rejection under § 103(a), per MPEP 2141 and *KSR*, the Examiner has the initial burden of making the findings necessary to support a prima facie case of obviousness. However, the Examiner has not pointed to any facts or reasoning that would support a conclusion that Gringeri's analysis underlying its buffer modeling and just-in-time scheduling was inherently flawed or limited with regard to its claimed advantages, such that adding on a layer of feedback control would provide any worthwhile improvement.

A further difficult issue, which, in the context of Gringeri, would apply equally to feedback or buffer modeling, is the issue noted above of what to do if a client-side underflow or overflow develops, in order to temporarily increase the transmission rate. Moreover, even if that new problem, inherent in the Examiner's suggested combination of Gringeri and Baker, could be solved (e.g., by some-

how quickly renegotiating the data transport contract, implementing it, and then readjusting the buffer model accordingly), the two-stage buffer-modeling approach of Gringeri already provides a viable implementation mechanism (in that context) for adjusting data flow during transmission, without the need for the considerable added complexity of attempting to graft feedback flow control on top of the systems already taught by Gringeri. The Examiner has hypothesized an advantage to be gained, but has not provided any reasoned basis for actually expecting such an advantage. In short, once the Examiner's suggestion is investigated in terms of possible implementation, any motivation for adding feedback control to Gringeri disappears. Such a lack of motivation again undermines the case for obviousness. See MPEP 2143(G).

For the reasons discussed, it also appears that adding Baker-style feedback to Gringeri on top of Gringeri's buffer-modeling methods would effectively change the principle of operation of Gringeri. Whereas Gringeri previously worked by blind server-side modeling, relying on the reliable nature of ATM transport and deriving transmission parameters that were bound to work successfully if properly monitored and applied during transmission, the proposed modification would forego reliance on the reliability of the ATM network, and instead change Gringeri to a system ultimately moderated based on real-time feedback flow control. However, if a proposed modification in accordance with a secondary reference would change the principle of operation of the primary reference -- as this modification would do -- then the combined references cannot be sufficient to render the claim prima facie obvious. MPEP 2143.01(IV).

In sum, because of (1) the inherent differences between ATM and IP protocols, (2) the complexity of any modification that would adequately work around these differences, (3) the lack of expectation of success for such workarounds, (4) the lack of motivation because there is no clear reason to expect that the proposed modification would provide an actual improvement, and (5) the fact that the proposed modification of the primary cited reference would change its principle of operation, Applicant submits the proposed modification of Gringeri in accordance with Baker would not be likely to provide a practical solution that was obvious to one of ordinary skill in the art, and fails to establish a *prima facie* case of obviousness.

Application of Applicant's Arguments to the Claims

The foregoing traversal of the present rejection of claim 7 applies with equal force to each of claims 8-14, which are each dependent claims the incorporate each and every limitation of claim 7. If the independent claim is nonobvious, "then any claim depending therefrom is nonobvious". MPEP 2143.03.

The same grounds of traversal apply to claims 15 and 16, which are method claims corresponding to claims 7 and 11.

In addition, there is a further ground of traversal with regard to claims 13 and 14. The Examiner rejected claims 13 and 14, based on the view that Gringeri, at Figure 4 and the corresponding text, discloses "pointer maintaining means for maintaining a record for each user computer of the last data element that had been sent by said server to said user computer, and for actuating said sending means when said user computer buffer is not full, to enable said sending means

to send additional data elements to said user computer at a rate as fast as the connection between said server and said user computer will support, said pointer maintaining means being arranged to maintain said pre-determined number of data elements in said user buffer", and "recording means for maintaining a record for each user computer of the last data element that had been received by said user computer, and, when said user computer buffer is not full, for requesting said sending means to send additional data elements to said user computer, said recording means arranged to maintain said pre-determined number of data elements in said user buffer, and said sending occurring at a rate as fast as the connection between said server and said user computer will support."

Applicant respectfully disagrees that Gringeri contains any such disclosures. Because of the deterministic nature of an ATM switched network, Gringeri presumes that all users of a given stream are at the exact same point in data receipt and playback, and therefore Gringeri requires no server-based process to keep track of the buffer status of the individual players (and does not provide any such facility). In particular, Gringeri discloses no means by which the server can know the state of any particular user device during the course of stream transmission, and no means for individualizing delivery to these devices. Figure 4 of Gringeri (which was cited in this regard by the Examiner) shows one (1) video decoder 34, and one (1) display 36. As discussed at col. 13, lines 15-20 of Gringeri, there may be more than one of these, and they can each independently request programming. However, there is no teaching or suggestion anywhere in Gringeri that any ***given*** video stream can be distributed differentially to a plurality of us-

ers, and no disclosure of mechanisms for individualizing distribution of a stream as recited in claims 13 and 14.

Accordingly, based on all of Applicant's foregoing arguments, Applicant respectfully requests reconsideration of the rejection under 35 USC § 103(a) of claims 7-16, and withdrawal of the rejection.

IV. Discussion of New Claims:

Applicant respectfully submits that new claims 11-59, added herein, define additional aspects of what applicant regards as his invention. Applicant maintains that features of applicant's invention discussed hereinabove in connection with the rejection of claims 7-16, whereby data is sent at the playback rate, and the user buffer is initially filled, and replenished at a higher rate, and in which data flow is controlled in accordance with feedback from the client-side buffer, are also present in claims 17-59. Thus, these claims are similarly not rendered obvious by the cited combination of Gringeri and Baker. As such, new claims 17-59 are submitted are submitted also to be in condition for allowance, and their allowance is respectfully requested.

CONCLUSION

In view of the foregoing, it is submitted that present claims 7-16, and new claims 17-59 are patentable over the cited prior art and are in allowable condition. Accordingly, allowance of claims 7-59 is earnestly solicited.

Respectfully submitted,
Harold Edward Price

By: ______________________

Ernest D. Buff
(His Attorney)
Reg. No. 25,833
(908) 901-0220